AN ANTHROPOMETRIC STUDY OF WOMEN WITH CANCER

DIANNA BRINKLEY, R. G. CARPENTER, AND J. L. HAYBITTLE

Department of Radiotherapeutics and the Medical Statistics Section, Department of Human Ecology,
University of Cambridge

Clinical observations over a long period suggested to one of us (D.B.) that many women with cancer of the breast were of a more masculine physical type than both normal healthy women and those with cancer of the cervix.

A number of published studies have also indicated that there may be some association between incidence of cancer and constitutional type. Sheldon, Hartl, and McDermott (1949) published some somatotype distributions of patients with cancer of the breast and uterus at the Presbyterian Hospital, New York, and the distributions showed the patients to be predominantly endomorphic mesomorphs. As this tendency also appeared in patients suffering from other diseases at the same hospital, it might well have been due to selection.

Hagnell (1961), by making personality assessments on a large number of Swedish women, found that a significantly high proportion of women who developed cancer were rated as substable personalities, i.e., warm, concrete, naive, industrious, sociable, and tending towards personal relationships, and Lindegard and Nyman (1956) showed that substability was correlated with body build.

Substability has traits in common with Eysenck’s (1959) ‘extraversion’, and Hagnell’s findings are therefore supported by the work of Coppen and Mactalfe (1963) who found a higher extraversion score in patients with breast cancer than in controls drawn from either a hospital or the general population. With these findings in mind, we began at Cambridge an anthropometric study of women patients with cancer of the breast and cervix to see if there were any physical differences between such women and a group drawn from the normal population. Measurements were also made on a group of hospital patients with cancer at sites other than the breast or the cervix and on a group of hospital patients without cancer.

It was hoped that from the information acquired a simple method might be evolved for detecting women with a ‘high risk’ of developing cancer of the breast. Such a method of detection could be of great value in a preliminary survey for breast cancer; elaborate tests could then be used in the high-risk group of women.

METHOD

The cancer patients were selected from those referred to the Radiotherapeutic Centre, Addenbrooke’s Hospital, the selection being determined by such factors as the patient’s fitness and willingness to co-operate, and their availability at the time when the measurements could be taken. Some were inpatients with new, untreated lesions, while others were attending the Centre for follow-up appointments. Patients both with and without recurrence were included in this latter group.

The group of hospital patients without cancer were all inpatients who were entered for the study through the collaboration of the consultant in charge of them. The only extra criterion used in their selection was that they should be in an age range comparable to that of the cancer patients, i.e., over 40 years of age.

The group drawn from the normal population was obtained through the co-operation of the Cambridgeshire Mass Miniature Radiography Unit. Volunteers for the study were asked for at a meeting held in the area before the visit of the MMRU to that area. Those volunteering were then interviewed and measured after having their chest x-ray taken. The number of subjects in each group is shown in Table I.

The anthropometric measurements taken were (a) weight, (b) height, (c) sitting height, (d) biacromial dimension, and (e) biliac dimension. The methods adopted for making these measurements were

<table>
<thead>
<tr>
<th>TABLE I</th>
</tr>
</thead>
<tbody>
<tr>
<td>NUMBER OF SUBJECTS IN EACH GROUP</td>
</tr>
<tr>
<td>Group</td>
</tr>
<tr>
<td>Cancer of cervix</td>
</tr>
<tr>
<td>Cancer of breast</td>
</tr>
<tr>
<td>Other cancer</td>
</tr>
<tr>
<td>Other patients with non-malignant disease</td>
</tr>
<tr>
<td>Normal women</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>
those recommended by Professor J. M. Tanner of the Institute of Child Health, (c), (d) and (e) being measured with special anthropometric calipers. Four other physical factors were also noted, namely hair colour, eye colour, hair texture, and the presence or absence of hirsuties. Hair colour was assessed as dark, medium, and fair. When the patient's hair was grey or dyed, she was asked to state its original colour. The assessment of hair texture (coarse or fine) and hirsuties (present or absent) was a very subjective one, but all the examinations and assessments were made by one observer (D.B.), so that it is hoped that the scale of assessment was reasonably consistent throughout the study.

Information on the following items was also obtained from each subject—age, marital state, menopausal state, patient's and husband's occupations, and the number of live births, still births, and miscarriages. From the occupation of the husband, or of the subject, if single, the social class category on the Registrar General's scale was determined. Complete data were available on all subjects included in this study. These were put on to punched cards. Before analysis started the data were pre-processed.

PRE-ANALYSIS DATA PROCESSING

The data on each patient were processed as follows. First, two new anthropometric measurements were calculated, the biacromial-biiliac ratio and the ratio of sitting height to total height. In view of the hypothesis that breast cancer patients tend to be more masculine than normal women, an androgeny index (Tanner, 1951) was also computed, i.e.,

\[
AI = 2 \times \text{biacromial dimension} + 0.53 \times \text{height—sitting height} - 1.25 \times \text{biiliac dimension}
\]

Hair colour was treated as a variable on a three-point scale, 1 = dark, 2 = medium, and 3 = fair. Three categories of eye colour were used, hazel or brown, blue or grey, and green. These were specified by two 1,0 variables; green, not green; and blue/grey, not blue/grey. Those with brown eyes are coded as being neither green nor blue/grey. The presence of hirsuties was recorded by a 1,0 variable and coarse hair texture was similarly coded.

Marital status was recorded in the same way as eye colour, i.e., by two 1,0 variables; single, not single; married, not married. Women who were widowed, divorced or separated were all classed as neither single nor married. The menopausal state was recorded by another 0,1 variable, 0 = post-menopausal; 1 = pre-menopausal.

Thus, after pre-processing, data for analysis comprise 19 variables. Eight are anthropometric, eight physiological, and three sociological. The revised data were put on to a second set of punched cards* and analysed by computer using the BMD Biomedical Computer Programmes (Dixon, 1966).

RESULTS

ONE VARIABLE AT A TIME

As a first step the five groups of patients were compared on each variable using a one-way analysis of variance. The computation was carried out with BMD 7D, which also gives the mean and standard deviations of each group, a plot of the data, and the correlation matrix for each group. The results are summarized in Figure 1.

Figure 1 shows the mean and confidence limits for the mean of each group of subjects for each of the 19 variables. Asterisks indicate whether a statistically significant difference was found in the analysis of variance. Confidence limits have been calculated from the pooled mean variance, except for binomial variables where they are based on tables (Hald, 1952) or charts (Pearson and Hartley, 1954). The following points may be noted from Figure 1:

- **% SINGLE** The percentage is highest, 18%, among the breast cancer cases, and lowest, 2%, among the cervix cancer group.
- **% MARRIED** The means for this variable present almost the reverse pattern to that of the previous variable. But the group with the largest percentage married is the Other Patients, which is also the youngest group.
- **SOCIAL CLASS** The low average social class of the cervix cancer group will be noted, and also the higher social class of the normals. These differences are reflected in the percentage of women in each of the five classes. From the 1961 census reports it can be calculated that in the Eastern Region the average social class of married women, classified by their husbands' occupation, is 3.07. The average social class of economically occupied women was slightly lower, 3.20, so that the average social class of the normal women, 80% of whom were married, is expected to be 3.10, which is significantly lower than that observed. The biggest difference is in the proportion of women in social class II, 33-6% observed compared with 18.7% expected. Three factors are probably collectively responsible for these differences—the area being X-rayed, over-reporting of the husband's occupation, and self-selection of those volunteering to be measured. By using social class as a co-variable in a multivariate analysis, as de

*Not all BMD programmes include routines or generating new variables.
scribed below, the effects of this self-selection can be at least partly eliminated.

**Age*** The average age of the groups varies substantially, the oldest being the breast cancer group. The spread of the ages in the five groups is, however, remarkably constant. The standard deviations of the groups are all within 10% of the pooled mean within-group S.D., 10.2 years.

**% Post-menopausal*** The pattern of the five means is almost the same as the pattern of the mean ages.

**Number of pregnancies*** The cervix cancer group had a large number of pregnancies. Only four of these women had never been pregnant and one had had 17 pregnancies, 14 of them live born. In contrast, 35% of the breast cancer patients had never been pregnant and only one had had seven children. All the distributions of the number of pregnancies are skew. The tail is especially long in the cervix cancer group. Also the standard deviations of the groups tend to increase with the group means.

**Sitting height** and **Height** Statistically highly significant differences between the groups were found in sitting height and height, and in the ratio of sitting height to total height. The Other Patients and normals were the tallest and have the largest sitting heights. The breast cancer group are shortest and have much the shortest sitting height. The result is that the average ratio of sitting height to total height is similar in all the groups except for patients with breast cancer, for whom the average value is significantly reduced.

**Biacromial and bililac** dimensions The pattern of the group means for both these dimensions is similar. The Other Patients have the largest average...
for both dimensions. The pattern of the mean is similar to that of height and sitting height and suggests that the Other Patients and normals tend to be bigger.

**Biacromial-Biiliac Ratio** This ratio shows no significant differences between the groups. In particular, there is no suggestion that the breast cancer group of patients tend to have masculine characteristics represented by relatively broad shoulders. If anything, it is the normal women who tend to be relatively broad-shouldered.

**Weight** Differences between the average weight of the groups are only just statistically significant. Inspection of the means shows that this is due to the weight loss of the Other Cancer group of patients. Otherwise there is very little difference between the weights in the five groups.

**Androgeny Index** No statistically significant differences emerge on this factor either. The mean score is lowest for women with cancer of the cervix and highest for Other Patients. There is no suggestion that the breast cancer group is particularly masculine as determined by the index.

**Hair Colour** and **Remaining Variables** No statistically significant differences were found on any of the remaining variables except for hair colour. This difference is due to the tendency for the breast cancer groups of women to have dark hair, 47% dark compared with 26% in the other four groups taken together. This may also explain the tendency for the breast cancer group to appear more hirsute.

**Multivariate Comparison of Groups**

The main differences between the groups can be seen from the foregoing analysis. However, the majority of the variables are to some extent correlated and some of them are highly correlated, e.g., age and menopausal state. In consequence, several variables may be providing the same information. Alternatively, some real differences may be concealed by the correlations. We required to know which of these variables contains statistically significant information for differentiating the groups that is contained in no other variable. Also we needed to compare the groups after taking account of differences in the age of the patients, i.e., we wished to be able to treat age as a covariable.

These problems were resolved by using the computer programme for stepwise multivariate discriminant analysis, BMD 7M (Dixon, 1966). A number of analyses were carried out so that no significant factors were missed. The accuracy of the computations was also checked by running the data* with both 7M and another programme, BMD 5M. Tanner's androgeny index was not included in this analysis. AI is a linear function of three of the other variables; therefore, if it were significant in combination with other factors, it would emerge spontaneously in the course of the analysis.

These analyses showed that 8 of the 18 variables were statistically significant. These were, in order of importance, age, number of pregnancies, biiliac dimension, weight, hair colour, social class, sitting height, and biacromial-biiliac ratio. Not all these variables are needed to discriminate between any pair or groups, but each variable contributes in some way to separating out one or another of the groups from the rest. With the aid of these eight variables each of the five groups may be shown to be statistically significantly different from every other at the 0-1% significance level. Table II shows how the patients are classified by these eight significant variables. In all, 48% of the subjects are correctly classified. If the allocation was purely random only 27% correct would be expected.

*Excluding 24 controls selected at random so that the data complied with the restrictions of BMD 5M.

### Table II

**Subjects Classified by Eight Statistically Significant Variables**

<table>
<thead>
<tr>
<th>Group</th>
<th>Percentage of Cases Classified into Group</th>
<th>No. of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cervix Cancer</td>
<td>Breast Cancer</td>
</tr>
<tr>
<td>Cervix cancer</td>
<td>38-3</td>
<td>15-0</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>7-3</td>
<td>64-0</td>
</tr>
<tr>
<td>Other cancer</td>
<td>24-0</td>
<td>18-0</td>
</tr>
<tr>
<td>Other patients</td>
<td>20-0</td>
<td>4-0</td>
</tr>
<tr>
<td>Normals</td>
<td>15-1</td>
<td>13-6</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>No. correctly classified</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>%</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table modified from the original.
Table III

### COEFFICIENTS OF STATISTICALLY SIGNIFICANT STANDARDIZED VARIABLES IN 10 PAIRWISE LINEAR DISCRIMINANT FUNCTIONS, STANDARD ERRORS, CONSTANT TERM, DISTANCE BETWEEN GROUPS AND EXPECTED PERCENT CORRECTLY CLASSIFIED

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coeff.</th>
<th>S.E.</th>
<th>Coeff.</th>
<th>S.E.</th>
<th>Coeff.</th>
<th>S.E.</th>
<th>Coeff.</th>
<th>S.E.</th>
<th>Coeff.</th>
<th>S.E.</th>
<th>Coeff.</th>
<th>S.E.</th>
<th>Coeff.</th>
<th>S.E.</th>
<th>Coeff.</th>
<th>S.E.</th>
<th>Coeff.</th>
<th>S.E.</th>
<th>Coeff.</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.73</td>
<td>0.18</td>
<td>0.26</td>
<td>0.15</td>
<td>-1.19</td>
<td>0.15</td>
<td>-0.50</td>
<td>0.21</td>
<td>-0.14</td>
<td>0.16</td>
<td>-0.08</td>
<td>0.16</td>
<td>0.08</td>
<td>0.14</td>
<td>0.06</td>
<td>0.13</td>
<td>0.06</td>
<td>0.14</td>
<td>0.05</td>
<td>0.20</td>
</tr>
<tr>
<td>Social Class</td>
<td>1.08</td>
<td>0.27</td>
<td>0.44</td>
<td>0.19</td>
<td>0.37</td>
<td>0.19</td>
<td>-0.14</td>
<td>0.28</td>
<td>0.11</td>
<td>0.26</td>
<td>-0.26</td>
<td>0.19</td>
<td>0.11</td>
<td>0.28</td>
<td>0.11</td>
<td>0.26</td>
<td>0.08</td>
<td>0.23</td>
<td>0.05</td>
<td>0.20</td>
</tr>
<tr>
<td>No. of Pregnancies</td>
<td>0.29</td>
<td>0.17</td>
<td>0.62</td>
<td>0.15</td>
<td>0.70</td>
<td>0.14</td>
<td>0.08</td>
<td>0.15</td>
<td>0.14</td>
<td>0.19</td>
<td>0.08</td>
<td>0.15</td>
<td>0.14</td>
<td>0.19</td>
<td>0.08</td>
<td>0.15</td>
<td>0.14</td>
<td>0.19</td>
<td>0.08</td>
<td>0.15</td>
</tr>
<tr>
<td>Biacromial/Biiliac Ratio</td>
<td>0.01</td>
<td>0.22</td>
<td>0.07</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
</tr>
<tr>
<td>Sitting Height</td>
<td>-0.26</td>
<td>0.37</td>
<td>0.39</td>
<td>0.39</td>
<td>-0.70</td>
<td>0.39</td>
<td>-0.14</td>
<td>0.39</td>
<td>-0.14</td>
<td>0.39</td>
<td>-0.14</td>
<td>0.39</td>
<td>-0.14</td>
<td>0.39</td>
<td>-0.14</td>
<td>0.39</td>
<td>-0.14</td>
<td>0.39</td>
<td>-0.14</td>
<td>0.39</td>
</tr>
<tr>
<td>Hair Colour</td>
<td>-0.37</td>
<td>0.37</td>
<td>-0.37</td>
<td>0.37</td>
<td>-0.37</td>
<td>0.37</td>
<td>-0.37</td>
<td>0.37</td>
<td>-0.37</td>
<td>0.37</td>
<td>-0.37</td>
<td>0.37</td>
<td>-0.37</td>
<td>0.37</td>
<td>-0.37</td>
<td>0.37</td>
<td>-0.37</td>
<td>0.37</td>
<td>-0.37</td>
<td>0.37</td>
</tr>
<tr>
<td>Biliac Dimension</td>
<td>0.01</td>
<td>0.22</td>
<td>0.01</td>
<td>0.22</td>
<td>0.01</td>
<td>0.22</td>
<td>0.01</td>
<td>0.22</td>
<td>0.01</td>
<td>0.22</td>
<td>0.01</td>
<td>0.22</td>
<td>0.01</td>
<td>0.22</td>
<td>0.01</td>
<td>0.22</td>
<td>0.01</td>
<td>0.22</td>
<td>0.01</td>
<td>0.22</td>
</tr>
<tr>
<td>Weight</td>
<td>0.22</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
</tr>
</tbody>
</table>

1) The S.D. here being the pooled mean within group S.D. of the discriminant scores
2) Modulus of coefficient > 1.96 S.E.
3) < 2.58 S.E.
4) > 3.29 S.E.

### Table IV

### COEFFICIENTS OF AGE IN PAIRWISE DISCRIMINANT ANALYSIS

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coeff.</th>
<th>S.E.</th>
<th>Coeff.</th>
<th>S.E.</th>
<th>Coeff.</th>
<th>S.E.</th>
<th>Coeff.</th>
<th>S.E.</th>
<th>Coeff.</th>
<th>S.E.</th>
<th>Coeff.</th>
<th>S.E.</th>
<th>Coeff.</th>
<th>S.E.</th>
<th>Coeff.</th>
<th>S.E.</th>
<th>Coeff.</th>
<th>S.E.</th>
<th>Coeff.</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.44</td>
<td>0.13</td>
<td>0.09</td>
<td>0.18</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>Social Class</td>
<td>0.07</td>
<td>0.16</td>
<td>0.33</td>
<td>0.16</td>
<td>0.28</td>
<td>0.16</td>
<td>0.28</td>
<td>0.16</td>
<td>0.28</td>
<td>0.16</td>
<td>0.28</td>
<td>0.16</td>
<td>0.28</td>
<td>0.16</td>
<td>0.28</td>
<td>0.16</td>
<td>0.28</td>
<td>0.16</td>
<td>0.28</td>
<td>0.16</td>
</tr>
<tr>
<td>No. of Pregnancies</td>
<td>0.12</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
</tr>
<tr>
<td>Biacromial/Biiliac Ratio</td>
<td>0.74</td>
<td>0.19</td>
<td>0.00</td>
<td>0.16</td>
<td>-0.60</td>
<td>0.16</td>
<td>-0.06</td>
<td>0.19</td>
<td>-0.22</td>
<td>0.19</td>
<td>-0.22</td>
<td>0.19</td>
<td>-0.22</td>
<td>0.19</td>
<td>-0.22</td>
<td>0.19</td>
<td>-0.22</td>
<td>0.19</td>
<td>-0.22</td>
<td>0.19</td>
</tr>
<tr>
<td>Sitting Height</td>
<td>0.10</td>
<td>0.28</td>
<td>0.10</td>
<td>0.28</td>
<td>0.10</td>
<td>0.28</td>
<td>0.10</td>
<td>0.28</td>
<td>0.10</td>
<td>0.28</td>
<td>0.10</td>
<td>0.28</td>
<td>0.10</td>
<td>0.28</td>
<td>0.10</td>
<td>0.28</td>
<td>0.10</td>
<td>0.28</td>
<td>0.10</td>
<td>0.28</td>
</tr>
<tr>
<td>Hair Colour</td>
<td>0.27</td>
<td>0.18</td>
<td>0.37</td>
<td>0.18</td>
<td>0.37</td>
<td>0.18</td>
<td>0.37</td>
<td>0.18</td>
<td>0.37</td>
<td>0.18</td>
<td>0.37</td>
<td>0.18</td>
<td>0.37</td>
<td>0.18</td>
<td>0.37</td>
<td>0.18</td>
<td>0.37</td>
<td>0.18</td>
<td>0.37</td>
<td>0.18</td>
</tr>
<tr>
<td>Biliac Dimension</td>
<td>0.01</td>
<td>0.22</td>
<td>0.01</td>
<td>0.22</td>
<td>0.01</td>
<td>0.22</td>
<td>0.01</td>
<td>0.22</td>
<td>0.01</td>
<td>0.22</td>
<td>0.01</td>
<td>0.22</td>
<td>0.01</td>
<td>0.22</td>
<td>0.01</td>
<td>0.22</td>
<td>0.01</td>
<td>0.22</td>
<td>0.01</td>
<td>0.22</td>
</tr>
<tr>
<td>Weight</td>
<td>0.22</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
</tr>
</tbody>
</table>

5.44, have a within-group standard deviation of the same.
5.45, within-group standard deviation is normal.
6.44, the corresponding percentages shown in Table III.
FIG. 2.—Distribution of age-adjusted discriminant scores for 10 pairwise comparison of groups.
Figure 2 (a–j) shows the distributions of the age-adjusted discriminant scores in within-group standard deviation units. The actual percentages of subjects in each group correctly classified by the corresponding discriminant function are also shown in Figure 2. These are generally fairly close to the predicted percentages shown in Table IV. The worst discrepancy occurs in the classification of the cervix cases (Fig. 2a–d) and is due to the skewness of the discriminant scores. This skewness stems from the marked skewness of the distribution of the number of pregnancies among these patients.

**DISCUSSION**

**Duplication of Information**

A study of the steps in the multifactorial analysis shows that age takes account of the significant differences in the menopausal state and, together with the number of pregnancies, accounts in a statistical sense for differences in the marital status of the subjects. The number of pregnancies is preferred to marital status as a factor because it discriminates between the groups more than does marital status, and at the same time incorporates together with age all the information provided by marital status.

The well-known tendency for women with cancer of the cervix to have had more children and come from a lower social class (higher numerically) emerges clearly in both the single factor and multifactor analyses. Similarly, women with cancer of the breast tend to have had fewer children than other groups, as is also well known.

Height is negatively associated with age (Miall, Ashcroft, Lovell and Moore, 1967), and differences between the groups in both height and sitting height are largely accounted for by differences in age, as the stepwise discriminant analysis shows. In contrast, taking out differences in age increases the significance of differences in the biiliac dimension and this is an important factor, as may be seen from the size of the discriminant coefficients in Table III—five of them for biiliac dimension are over one. Weight is also important.

After six steps of the stepwise process the anthropometric measurements, biacromial-biiliac ratio, ratio of sitting height to total height, sitting height, height, and biiliac dimension, are all marginally significant variables that have not been included. At the seventh step the ratio of sitting height to total height is tried but this variable is later rejected in favour of a combination of the biacromial-biiliac ratio plus sitting height. In fact, there is probably little to choose between any of these five variables and no special significance should be attached to the particular combination selected.

**Anthropometric Discriminant Functions**

If one of the discriminant functions is applied to two groups of women, who are all the same in respect of one of the variables, the variable contributes nothing to the discriminant and may be omitted. Coefficients of other variables are unaffected.

Thus, if a group of women with breast cancer were compared with normals, all having the same age, social class, number of pregnancies, and hair colour, the discriminant function would involve only the remaining anthropometric variables, biacromial-biiliac ratio, sitting height, biiliac dimension, and weight. The estimated coefficients of these variables would be as shown in Table III, i.e., 0.59, -0.67, 0.31, and 0.15 respectively. When the insignificant variable weight is excluded, we find an anthropometric index for discriminating between women with breast cancer and normals all having the same age, social class, number of pregnancies, and hair colour is*

\[
I(B,N) = 0.62 \text{ (biacromial-biiliac ratio)} + 0.42 \text{ (biiliac dimension)} - 0.65 \text{ (sitting height)}
\]

Also a comparison of otherwise comparable women with breast cancer and cancer of the cervix gives a very similar index†

\[
I(B,C) = 0.54 \text{ (biacromial-biiliac ratio)} + 0.40 \text{ (biiliac dimension)} - 0.39 \text{ (sitting height)}
\]

If the measurements used in Tanner's index of androgeny are in standardized units the coefficients become

\[
AI(\text{standardized measurements}) = 4.0 \text{ (biacromial-dimension)} - 2.1 \text{ (biiliac dimension)} + 3.4 \text{ (total height)} - 1.9 \text{ (sitting height)}
\]

Now the mean of the standardized biacromial dimensions is 17.7 units and that of the standardized biiliac dimensions is 17.6 units and their correlation is close to 0.5. So that for small deviations from the mean, the standardized biacromial-biiliac ratio is close to the difference between these two dimensions. Hence I(B,N) and I(B,C) may approximately be re-written:

\[
I(B,N) = +0.62 \text{ (biacromial dimension)} - 0.20 \text{ (biiliac dimension)} - 0.65 \text{ (sitting height)}
\]

*When weight, an insignificant variable, is excluded from the discriminant function, biiliac dimension becomes significant in I(B,N) and is very nearly significant in I(B,C) and has therefore been included in both. The coefficients of the other components are also slightly modified by the exclusion of weight.

†Changing the signs so that breast cancer patients score positive with both indices.
I(B,C) = 0.54 (biacromial dimension)
- 0.14 (biliac dimension)
- 0.39 (sitting height).

It therefore appears that apart from an arbitrary scale factor these two indices are essentially similar to Tanner's androgeny index. The most important difference is that in Tanner's index total height has been included. Men and women differ much more in height than do these three groups of women, which probably explains the inclusion of this variable in Tanner's index. Thus from our data it appears that, other things being equal, women with breast cancer have a larger biacromial-biliac ratio for a given sitting height than normal women or women with cancer of the cervix, i.e., are more masculine as judged by an androgeny index similar to Tanner's.

In other contrasts in Table III the coefficients of the biacromial-biliac ratio and sitting height are only significant twice and in these other anthropometric variables are involved besides.

**EFFECT OF BREAST CANCER TREATMENT ON THE ANTHROPOMETRIC INDICES**

Both anthropological indices, I(B,N) and I(B,C), suggest that women with breast cancer tend to have relatively broad shoulders for their sitting height and hip width. These indices would be inflated if breast cancer or its treatment caused these women to become hunched up as this would reduce the sitting height (see Fig. 1). These factors can be fully studied only by a prospective study. However, some indication of the effects of these factors on the indices may be obtained from a comparison of the three groups of women in the breast cancer group. These are (1) new untreated patients, (2) previously treated patients without recurrence, and (3) previously treated patients with recurrence. Treatment included both radiotherapy and/or surgery.

The three groups of patients were first compared using the eight variables that were significant in the overall discriminant analysis (Table III). No statistically significant differences were found between the three groups of breast cancer patients on these variables, nor are any of the mean differences at all large. Second, after adjusting the observations on each patient for differences in age, social class, number of pregnancies, and hair colour, the index I(B,N) was evaluated. Figure 3 shows the mean value of this standardized index score for the three groups of breast cancer patients and for the normals.

As previously shown, the discriminant function I(B,N) gives a statistically very highly significant difference between the breast cancer group and the normal women. Figure 3 shows that the confidence interval for the mean score for normal women do not overlap the confidence intervals shown for any of the other three means. But Fig. 3 also shows that the mean scores of the three groups of breast cancer patients vary, and analysis of variance shows that the differences between the three means are statistically significant (0.05 > p > 0.01). The largest difference is between the untreated patients and those who had been treated without recurrence.

Thus these data suggest that the breast cancer or its recurrence inflates the mean value of the index I(B,N), and that successful treatment tends to reduce the mean index score towards normal. These differences may, however, be largely due to chance. The discriminant functions I(B,N) and I(B,C) are so similar that it is pointless to repeat the analysis with the latter function.

**CONTRIBUTION OF ANTHROPOMETRIC DATA TO DISCRIMINATION**

The contribution of the anthropometric measurements to the discriminant functions has been studied in detail for the comparison of the two groups, breast cancer and normals. Carpenter (1971) shows that this may be done by analysis comparable to that of making age a covariable. Such an analysis shows that given women from these two groups having the same age, social class, and number of pregnancies, a discriminant function, using hair colour and the anthropometric variables in Table III, would expect to classify 66.4% of patients correctly compared with 50% if allocation was by chance. If the women had the same hair
colour the anthropometric measurements alone would expect to classify 61.8% correctly. But given women from these two groups of the same age, one might expect to classify 61.8% correctly on the basis of social class and the number of pregnancies, and, by using hair colour and the anthropometric measurements, the expected percentage correctly classified increases only 8.1% to 69.9%. Thus although the anthropometric measurements are statistically highly significant in the percentage of subjects correctly classified that is obtained from them is relatively small. These percentages are also based on the assumption that the same proportion of errors is made in both groups.

To use a discriminant function as a means of screening women to find those with a high risk of having breast cancer it would be important to choose the discriminant point so that no more than 5% of women with breast cancer were excluded. But in this case, using the discriminant function of Tables III and IV, only 29% of normal women of an age to be at risk of developing breast cancer would be excluded by the screening test, i.e., 71% of the normal women would still have to undergo further examination. This can be seen very clearly from Figure 2g. Thus for practical purposes the discriminant functions described have relatively little value. The best we could hope for is that these anthropometric differences may provide some extra information which will improve the discriminating power of some other test, e.g., one based on steroid excretion (Hayward, Bulbrook, and Greenwood, 1961), which Wade and his colleagues have recently shown to be unsuitable for diagnostic screening (Wade, Davis, Tweedie, Clarke, and Haggart, 1969).

SUMMARY

An anthropometric study has been made of female patients with cancer of the breast and cervix to see if there are any physical differences between such women and a group drawn from the normal population. Measurements were also made on a group of hospital patients with cancer at sites other than the breast or the cervix, and on a group of hospital patients without cancer. A total of 509 subjects were studied. After preprocessing the data, 19 variables on each subject were available for analysis: eight were anthropometric, eight physiological (e.g., eye and hair colour), and three sociological (e.g., marital state and social class).

A stepwise multivariate discriminant analysis showed that eight variables were statistically significant. These were, in order of importance, age, number of pregnancies, biiliac dimension, weight, hair colour, social class, sitting height, and biacromial-biiliac ratio. Anthropometric indices were calculated for discriminating women with breast cancer from normals of the same age, social class, number of pregnancies, and hair colour, and also from comparable women with cervix cancer. These indices are very similar to Tanner's androgeny index, and suggest that women with breast cancer tend to be more masculine in type.

However, although the anthropometric measurements are statistically highly significant in distinguishing breast cancer patients, a discriminant function using these variables would still be very inefficient as a method of screening women to find those with a high risk of having breast cancer.

We are grateful to Professor J. M. Tanner of the Institute of Child Health for advice on making anthropometric measurements. We wish to thank Dr. Daphne Smith and the staff of the Cambridgeshire Mass Miniature Radiography Unit for the help they so willingly gave in the measurement of normal subjects, and also the many women whose co-operation made this study possible. We are greatly indebted to Professor Mitchell for his interest and support throughout the project, and to Imperial College, London, and the Cambridge University Computer Laboratory for access to their computers.

REFERENCES


ANTHROPOMETRIC STUDY OF WOMEN WITH CANCER


APPENDIX

Pooled mean within-group standard deviations used to standardize variables used in the discriminant functions:

<table>
<thead>
<tr>
<th>Variable</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>10:24</td>
</tr>
<tr>
<td>Social class (5 pt. scale)</td>
<td>0:953</td>
</tr>
<tr>
<td>No. of pregnancies</td>
<td>1:92</td>
</tr>
<tr>
<td>Biacromial/biiliac ratio</td>
<td>0:0757</td>
</tr>
<tr>
<td>Sitting height (cm)</td>
<td>3:49</td>
</tr>
<tr>
<td>Hair colour (3 pt. scale)</td>
<td>0:616</td>
</tr>
<tr>
<td>Biiliac dimension (cm)</td>
<td>1:66</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>11:44</td>
</tr>
</tbody>
</table>